Evaluation of the Validity of Medical Records in Trauma Patients at the Pine Ridge Hospital. Christopher Krogh, Class of 1992.

The medical record has many uses beyond that of patient care. Additional uses include epidemiologic studies and other forms of research, quality assurance, public health policy, health-care financing, billing, and fiscal management. Without valid information, such activities build upon a questionable foundation; and in a "worst-case" scenario, good care or positive fiscal performance may be undermined because of bad data.

The International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9CM) is the coding system most commonly used in US medical institutions to record clinical diagnoses and some procedures. ICD-9 codes are clinical and presuppose that one has been able to achieve a diagnosis; they are less suited to ill-defined or vague conditions such as sore throat with negative culture, abdominal pain without other clinical findings, or multiple small injuries. There are no codes at all for many unconfirmed or "rule-out" conditions. On the other hand, there are many codes for other conditions. For example, a recent edition of the ICD-9 manual contained 146 different code numbers for the variants of middle ear infection (otitis media).

STATEMENT OF PROBLEM

E-codes (a type of ICD-9 code) are generated as an outcome of injury-related hospital visits. These codes were developed specifically to reflect the forces, effects, and circumstances involved in injuries. Their validity is therefore crucial in the evolving field of injury control. If a significant proportion of E-codes are invalid, then many conclusions which are being reached about the causes, nature, outcomes and costs of different injuries may likewise be suspect.

Since injuries are straightforward and easily documented, one might expect E-codes to be some of the most accurate of all ICD-9 codes. Even in the case of E-codes, however, there is increasing evidence that they do not always correspond to injuries as they actually occurred. Previous investigators have compared E-codes to hospital records with varied findings. Moreover, one may question how well hospital records themselves actually describe the injury as it happened. A variety of factors - such as an intoxicated or unconscious patient, a large number of wounds, a very busy clinician, as well as legal and political concerns, may render it difficult for a clinician to accurately describe the nature and circumstances of the injury.

During the summer of 1992, a medical student became available for three months to independently observe the presentation of injuries and their management at one Indian Health Service emergency room and to compare his independent observations with hospital records and the resultant E-codes. This provided an outstanding opportunity to clarify if, and to what extent, E-codes and records reflect the independently-observed reality.

QUESTION AND HYPOTHESES

How often, and to what degree, do E-codes and hospital records resulting from injury-related encounters actually reflect what happened to the patient?

Hypothesis 1: There is no difference between information obtained from medical records and E-codes and that obtained by direct observation of trauma visits.

Hypothesis 2: There are discrepancies between information obtained from medical records and E-codes and that obtained through direct observation, but the discrepancies are systematic and predictable and, therefore, may be corrected.

METHODS

The study was conducted from September 10 to October 20, 1992. Working in shifts which included all hours of the day, fourth-year medical student James Clover observed in the Pine Ridge Emergency Room as trauma patients were examined and treated. He kept independent records as to the nature, location and extent of all injuries as well as whether the patient, or another individual involved in the incident, appeared to have (or admitted, or was found on blood test to have) been drinking immediately prior to the accident.

Subsequently, Clover looked up the ICD-9 and E-codes which corresponded with each injury case he saw. Later, and independently, he compared his notes and codes with the hospital emergency room records and with the ICD-9 codes and E-codes that had been assigned by the hospital to each visit.

RESULTS

The Sample Population

The study sample consisted of 108 people who presented for treatment at the Pine Ridge Emergency Room between September 11 and October 20, 1992. In five cases, a chart could not later be located, resulting in a working sample of 103. These 103 did not represent every patient who presented to the ER during the study period. Periods of observation included all 24 hours of the day and all 7 days of the week. However, the observer was present in shifts (not continuously), and it was not possible to observe multiple simultaneous encounters.

   72% of patients gave their own information and histories. 21% of information came from a parent or guardian. 6% of information came from ambulance staff. In one case, the history was provided by a police officer.

   The trauma included 34% falls, 15% motor vehicle accidents, 13% assaults, 12% lacerations, and 26% other.

   In 32% of cases, alcohol was involved. In all but two of these it was the injured person who had been drinking. Only 48%, (16) of these were recorded on the ER sheet as alcohol related. 11 of the victims were coded by the hospital as intoxicated; this is a subjective judgement, but student Clover independently noted an additional 19 should have been so coded.

The Coding Process

   The medical record includes blanks for "purpose of visit", usually filled out by a physician; and "cause of injury" and "place of injury", usually filled out by a nurse. A coding clerk in the medical records department uses these three pieces of information to generate an E-code. If insufficient information is present, the coding clerk will examine the nurse's notes and possibly the physician's notes for additional information.

Agreement between observations, records, and codes

The first and (where recorded) the second, third and fourth diagnoses in the medical record were identified and we determined whether each diagnosis was supported by (I) direct observation or (2) the e-codes generated by the hospital.

AGREEMENT BETWEEN: MEDICAL RECORD				
	1ST DX	2ND DX	3RD DX	4TH DX
OBSERVATION	64%	37%	44%	33%
E-CODES	73%	58%	60%	25%

The E-codes generated by the hospital for the first and (where they existed) the second, third and fourth diagnoses were examined to determine whether direct observation had supported them.

E-Codes				
	1st Dx	2nd Dx	3rd Dx	4th Dx
Observation	44%	27%	28%	8%

All E-codes from these visits (not distinguishing between the first, second, third and fourth diagnosis) were examined to determine whether they were supported by (1) direct observation and (2) the medical records.

   This comparison differs from the previous ones in that (1) this comparison of observation and E-codes includes E-codes for all four (or more) diagnoses as a group; and (2) deciding whether E-codes are supported by observations is different from determining if observations are reflected in E-codes. For example, an E-code may stand for superficial abrasion (scratch) of the face. This might be borne out by actual observation of a crushed skull, since undoubtedly facial abrasions would be there too. But the reverse is not true: a crushed skull and head (for which very specific E-codes are available) is not appropriately coded as a facial scratch.

AGREEMENT BETWEEN:	E-Codes	Medical Record
OBSERVED	28%	77%
MEDICAL RECORD	42%	

Types of Discrepancy

(1) The commonest source of discrepancy was use of very nonspecific codes (such as E 928.9 - unspecified accident, or E819.1 - MVA of unspecified nature, unspecified person) when much more specific or appropriate information could have been coded (such as E816.1-loss of control by driver on highway, without collision). This often occurred if the physician or nurse entered inadequate or illegible information, and in some cases it appeared that details were not sought. For example, a patient might be noted to have fallen without determining if it was a fall down a stairs or a fall from playground equipment. Another reason for this problem is that coders might memorize several common or nonspecific codes and then prefer to use those rather than look up more detailed codes in the huge, and difficult to use, ICD-9 coding book.

(2) Apparent misreading or misinterpretation of information. As an example, "neck sprain" (847.0) was coded as "neck pain" (723.1, a different series altogether); "wrist strain" (842.00) was coded as "unspecified derangement of joint, hand" (718.94). Sometimes codes that have no logic at all are produced: "contusion to the forearm" (923.11) was coded as "neuralgia, neuritis, radiculitis unspecified" (729.2), for example.

(3) Some coders appear to follow informal rules which differ for different coders. Use of nonspecific codes is sometimes one manifestation of this. As another example, a person thrown from a horse (E 828) may be coded as E 906.9, "other injury caused by an animal", which is figuratively but not literally correct.

(4) Diagnoses or anatomical sites may not be correctly recorded. As an obvious example, where does the "elbow" end and the "forearm" begin?

 (5) Multiple injuries are not always documented in cases where some are major and others appear minor.

<u>DISCUSSION</u>

   The poor agreement between direct observation, medical records, and E-codes has led us to reject both hypotheses of this study - the first for reasons already presented, and the second through examinations of patterns in the data which are not presented in this short paper.
   Our finding that information in the E-codes is not supported by actual observation in more than 70% of cases suggests that E-codes (and, in fact, hospital emergency room records), represent a mythology rather than an accurate indication of what actually happened. The reverse finding--that what actually happened is not reflected in E-codes in 62% of cases--suggests that additional valuable information, and possibly reimbursable services, are being missed. If more than 70% of E-codes do not reflect what actually happened and if, in addition, two-thirds of what actually happened does not appear in E-codes, it seems clear that E-codes cannot support research, prevalence studies, quality assurance, billing, or any other function which requires data that is accurate at least half the time.

<u>RECOMMENDATIONS</u>

(1) A lack of familiarity with medicine and medical terminology makes it difficult for a coding clerk to interpret, and code, the information in a medical record. Coding staff could be offered basic training in anatomy and medical terminology.

(2) Some physicians and nurses do not record information necessary for reconstructing the patient's injuries or for generating E-codes. This could be corrected by legibly writing all relevant information under "purpose of visit".

(3) Physicians and nurses should have some basic information about ICD-9 codes so that they will understand what must be documented to assure that a correct code can be generated. At a minimum, they could be shown how codes are organized in a code book and how much information is required for an accurate code.

(4) More standardized forms, perhaps Severe Injury Surveillance Data Forms, should be available to the staff in each Emergency room. This would assure a more systematic, shared approach to injury recording which in turn could help to optimize injury prevention and surveillance efforts as well as care.

<u>CONCLUSIONS</u>

It could be argued that this project involves the data from a single hospital which does not represent "the norm" of IHS. Other investigators are challenged, therefore, to determine whether these rather alarming findings can be reproduced in other hospitals. A growing literature suggests that this is not a hospital-specific problem and that fundamental characteristics of medical information, and how it is acquired and coded, will similarly affect E-codes across clinical settings.

The use of a medical student from another country as the "standard" of what actually occurred is also perhaps open to question. Two arguments, however, support this strategy. First, a student, in the clinical years, from England comes with no preconceptions of US or Indian Health care and is arguably a more objective observer in this setting than a US student or an IHS investigator who may have many preconceptions. Second, injuries are injuries; they are straightforward to describe and a 6-cm laceration of the forearm, for example, does not require cultural sensitivity, insight, or clinical logic to identify and characterize. Therefore, our student (who, incidentally, did have these qualities, whatever their relevance) could be expected to be a reliable observer in the case of injuries as well as being an objective one.

The reader may also question how serious the discrepancies actually are. After all, characterizing a "cold" as an "upper respiratory infection" is a discrepancy, but the basic meaning is the same. How serious is it when ICD-9 codes do not match? In part, the answer depends on the very nature of ICD-9 codes and the uses made of them. ICD-9 codes are extremely specific; and for such codes to work, they must be chosen appropriately for the information available. If a vague code is used when specific information is known--or a very specific code used when available information is vague--then ICD-9 codes exaggerate the shortcomings of available information. Moreover, researchers conducting prevalence studies from ICD-9 codes, or hospitals conducting QA reviews, in most cases, do so with computers which cannot identify which codes "almost" match and which are widely discrepant, or decide how serious those discrepancies are. Thus either two codes match or they do not; the magnitude of discrepancy is a secondary point. The difference between "abdominal pain" and "appendicitis" may be much more serious than that between "throat pain" and "upper respiratory infection", but the error--the difference between a specific and a general code--is the same.

If hospital records and ICD-9 codes related to injury--one of the most straightforward and readily-observed forms of medical complaint--do not accurately reflect what was seen more than two-thirds of the time, then it appears very likely that records and codes for other, less clear-cut conditions are even less reliable. Thus, work in the relatively clear-cut domain of injury surveillance may pave the way for crucially needed reexamination of clinically-generated data.